Condensed Guide to the Binet Tests

By S. D. Porteus, Director of Research, Training School, Vineland, N. J.

Part I

Binet Tests and Diagnosis

It is now about fifteen years since the first publication of the already famous Binet-Simon Tests. Though this would be normally too short a time to test their real value, still experience has been so multiplied in their use, that we are at least beginning to get just views of their limitations and value. In some respects the tests have failed to fulfil expectations founded on them. They do not constitute a perfect instrument of research, as they cannot be relied upon for the accurate diagnosis of the highest grades of feeble-mindedness. This does not mean that their estimates of test age are erroneous. The difficulty is that there are factors besides the degree of intellectual development which have a bearing on social competency. In other words, there are persons intellectually inferior who are normal by the social criterion, because they possess practical abilities not evaluated by the Binet. On the other hand, there are individuals who are intellectually normal, but who are socially unfit, because of instability or weakness of temperament or volition, faults which are not revealed by Binet examinations.

The proofs of the validity of the Binet for diagnostic purposes have been mainly based on mass results, and this agreement with experience in a very large number of cases has been sometimes taken as sufficient proof of universal validity. This is unjustified. Diagnosis is an individual matter, and correct diagnosis is just as important in the case in which intellectual capacity is not the main factor in social fitness as it is in the case in which a Binet test age *does* give a fair indication of social efficiency. Considering the possibly tremendous importance to the individual of correct diagnosis, the examiner cannot afford to make mistakes. A method which is a reliable guide for even eighty per cent. of cases must always require corroboration, unless, of course, the method itself distinguishes the cases in which its verdict is of diagnostic significance

from those in which the intellectual inferiority is not the essential point. Unfortunately the Binet cannot do this. Any Binet limits of mental abnormality high enough to include all defectives would include also many normals. Any limits low enough to exclude all normals would exclude also many defectives. Hence it follows that we cannot with justice regard a Binet test age as conclusive proof of social unfitness in an individual.

But not only must a comparatively low Binet age be regarded as insufficient proof of potential social incompetency, but a high Binet intelligence quotient must not be regarded as irrefutable evidence of superior intelligence, using intelligence in the sense of capacity of the highest social or community value. Here again it is rather interesting to see the tendency to look upon results which are true in the main as being of universal application. This is exemplified in Terman's* treatment of the subject of the relationship between Binet test age and super-intelligence. Whilst ready enough in many instances to invoke teachers' judgments as proof of the validity of the tests, he attaches but little weight to the negative instances. Unfavorable reports as to moral characteristics were obtained from teachers with regard to eight out of thirty-one children with very superior Binet records. Two were "very self-willed," one needed "close watching," one was "cruel to animals," one was "untruthful," one "unreliable," one a "bluffer," and one "sexually abnormal, perverted, and vicious." These characteristics are put aside as not being "essentially abnormal," the inference being that the teachers' descriptions were merely overstatements.

A recent study by Anna Gillingham of children with very high intelligence quotients revealed the fact that many of those who were full of promise were showing remarkably poor scholastic performance. Some of these will undoubtedly, in adult life, belong to the class of brilliant failures with whom we are all familiar. It is commonly said of these, that they have brains, but lack application, or that they have such a conceit of their own abilities that they will not accept the drudgery essential to success. In a long clinical experience, the writer has met with several cases with I. Q.'s so high that Terman would place them in the "near genius" class, but who were already socially so ill-adapted as to require as much protection for their own good or for the good of others as any feeble-minded child. It is noteworthy that in each case referred to, tests of common prudence and foresight placed the children from eight to nine years lower in test age than did the Binet. There is a tendency to conceal a lack of understanding of such cases by applying to them the term "psychopathic," even when no psychosis is demonstrable.

The lack of relation between actual social worth and the Binet test ages of some of these individuals is due to the fact that the tests favor certain

*See "Measurement of Intelligence," p. 15.

371.92
T68
v.17-18

types of ability. An analysis of the character of the individual tests shows the following facts:—No less than fifty-seven questions out of seventy-four require an oral response, whilst language development, either from the standpoint of comprehension, range of vocabulary, descriptive or defining power, is the main capacity tested in fifty per cent. of the tests. Tests very largely dependent on previous educational training make up thirty-one per cent. of the total. Immediate memory is tested directly in twenty-one per cent. of the tests and in fifty per cent. of the alternative tests, whilst mediate or immediate memory is one of the main factors in success in no less than forty-eight per cent. of the total tests. What we may designate as tests of practical ability, requiring a motor response, constitute only seventeen per cent., whilst arithmetical ability, with twelve per cent., has almost as much attention. The criticism that may justly be founded on this analysis is that the tests are far too "literary." They favor the glib-tongued, quick-thinking child, the child who has had a good educational environment, the child who memorizes readily and to whom expression is easy, and who, therefore, shows good scholastic promise. On the other hand, they do not do justice to the child with capacity for achievement in manual and industrial pursuits, who is shy and unresponsive in oral tests, or who lacks facility in language expression, or whose mental operations are slow but sure, or who has been unfortunate in his educational background. In brief, while designed to test general ability, they do not succeed in being sufficiently general to suit all types of intelligence.

Naturally the type of child who is favored by the tests is most likely to show school proficiency. Hence, it is not surprising that the Binet tests are proved to have a very high correlation with educational capacity.* Indeed, it is mainly through comparison with teachers' judgments that the validity of the tests has been established. On the other hand, it must be admitted that there are children with high Binet records who show a faulty power of adaptation to school conditions. This is sometimes, but by no means always, the fault of the school. The class dullard, too, sometimes makes good. In either case, a Binet record may be a one-sided view. Provided the limitations of the tests are recognized, there is very little danger of false conclusions, and since there are varied types of ability, it is not necessary to attempt to stretch one series of tests so thin as to cover the whole ground. The need is not so much to reform the Binet as to balance its verdicts against other properly standardized, supplementary tests.

It must be emphasized that this discussion of the limitations of the Binet applies only against its exclusive use as a means of diagnosis. The recent army examinations of recruits have probably done more than anything else to help us realize that a mental age rating by any single series of tests is utterly unreliable as a measure of social competency. A very large percentage

*For correlation tables see "Porteus Tests—Vineland Revision," pp. 19, 20.

206980 59

of men, presumably socially well-adapted, was found to be distinctly below the levels of test age that were at one time thought to constitute the lowest limits of normality. Where two series of tests are used and different results accrue, we realize that the differences do not mean mutual contradiction, but that both tests may be right because neither of them are general enough to make a comprehensive test of abilities necessary for social fitness. Each of them is reflecting a different facet of intelligence. With regard to the Binet it may be said that, despite all its limitations for complete diagnosis, it will continue to play the most essential part in mental examination. The value of the tests is fully recognized when we say that they constitute, by far, the greatest contribution of psychology to applied science.

The first revision, and for long the most popular form of the tests, was that by Dr. Goddard. To his experimental work with the tests belongs the honor of the rapid recognition in this country and in many other countries as well, of their value as aids to mental diagnosis. Dr. Goddard's restandardization of the tests, though imperfect in some respects, made them far more serviceable. Binet's work, though evidencing genius, was carried out with comparatively few subjects, and it is extremely doubtful whether the extension of their use would have reached to a fraction of its present degree if it had not been for the confirmatory work and the publicity given to the results by Dr. Goddard. Very many thousands of children have been examined by Goddard's revision. Then followed revisions by several workers, none of which has attained as wide-spread recognition. The Yerkes-Bridges "Point Scale" varied the arrangement of the scoring system of the tests, added new questions, and in several respects made a distinct advance. The Stanford revision by Terman and others has, of recent years, made great headway. It attempted to correct errors in standardization, which were felt to be present in the Goddard Revision. Generally summarized, the Terman Revision differed from the Goddard in an increase of the difficulty of the tests for lower years and in a lessening of difficulty of those for the upper years, the provision of an equal number of tests for each year (except Year XII), a more refined method of scoring, and the needful extension of the Scale for adolescent years.

As regards the altered standardization, the changes were quite justified, as will be shown in a succeeding table. The chief objection to the Stanford Revision is that it has been standardized by application to children of middle class, who are considerably above the social and intellectual average of the great mass of delinquent or defective children to whom the scale is usually applied. This standardization of the tests by results from a socially selected group has led to erroneous conclusions as to the average level of test-age development in the general population, with consequent serious errors in individual diagnosis. Terman concluded that the average level should be put

at sixteen years, whereas the recent work of the army examiners has shown that it is much below this level. Consequently, it was possible to diagnose an adolescent as being at feeble-minded levels, when, as a matter of fact, he was comparatively little below the average of the ordinary population.

The question of correct standardization of the tests will never be finally settled until we have clear ideas as to the effect of selection of cases from various social groups. Each restandardization of the Binet that finds favor means the discarding of thousands of individual results obtained with previous revisions, which might otherwise be available for research. This has been the case with the Goddard Revision. Scores of thousands of test results are at present useless for comparative purposes, because of the widespread adoption of later revisions. Table I presents evidence pointing to the fact that the Terman Revision is still too easy in the lower years, and too difficult in the upper years for the average school child. The reason that this was not apparent in Terman's group of cases is probably due solely to the fact that the cases represented by Table I were of different social grade from those examined by Terman. The difference between average chronological age and test age at some age levels are shown to be so marked as to indicate that a further restandardization is necessary, with the probable effect of adding to the already existing confusion. A measuring scale whose divisional units are constantly changing is not very satisfactory as an instrument of research.

TABLE NO. I.

BINET-SIMON TESTS. STANFORD REVISION.

MENTAL AND CHRONOLOGICAL AGES COMPARED

1000 Cases—(Porteus)

Chron. Age.	No. of Cases	Average Mental Age	Difference Mental & Chron. Age
5.5 yrs.	30	6.24 yrs.	+ .74 yrs.
6.5 "	71	7.3 "	+ .8 "
7.4 "	124	7.7 "	+ .3 "
8.4 "	152	8.5 "	+ .1 "
9.4 "	140	9.1 "	— .3 "
10.4 "	119	10 "	— .4 "
11.5 "	114	10.5 "	—1.0 "
12.5 "	127	11.3 "	—1.2 "
13.4 "	98	12.8 "	— .8 "
14.3 "	25	12.5 "	—1.8 "

We do not think that it has been sufficiently recognized that the percentile method provides a remedy for erroneous standardization, and at the same time gives us a basis of comparison between the results by different revisions of the tests. Many writers have urged the advisability of arranging all mental measurements in the form of percentile tables, and several investigators have adopted their suggestion. Noteworthy amongst these are Pintner and Paterson, who have published percentile tables for their "Performance Scale." Woolley, also, in her "Scale of Mental and Physical Measurements for Adolescents," has used the method with evident advantage.

Considering the practical value of percentile tables and the ease with which, by their means, results of dissimilar tests may be compared, it is rather remarkable that more percentile tables have not been published for other scales and particularly for the Binet.

The difficulty appears to be that there have been no sufficiently large groups of normal children examined, and in order to compare percentiles by different tests, a large number of cases must be available and there must be approximately an equal distribution of ability. When this is the case, and the percentile tables have been arranged for the various social grades, errors in standardization will be automatically corrected. For instance, if by the Goddard Revision of the Binet six-year children are given a median test age of seven years, then seven years would equal the fifty percentile of the tables. Supposing the Stanford Revision placed the same children at a median level of five years, test age, then this age would equal the fifty percentile of the group. Comparing these children by test ages, the Goddard Revision would place them two years higher in mental age than the Stanford. Comparing percentiles, however, both series of tests would place the children on the same mark. Thus, provision of these percentile tables would mean that a vast amount of discarded data would continue to be available for research. As regards the tests themselves, we could be quite satisfied with a reasonable correctness of standardization, instead of an impossible accuracy. Of course, different social grades would need their own percentile tables, which would mean the further examination of large groups of children, but when these had been provided, the necessity to tinker with the tests would be quite done away with. Further, the means of comparison between mental tests and various physical measurements, which are already in percentile tables, would be measurably improved. In addition, percentiles are so easily capable of interpretation, that they can be easily understood by any person without special scientific training.

Realizing these advantages, we are publishing herewith, percentile tables of Binet mental measurements, based on a large number of cases. Table II gives the results for three thousand cases examined by the Goddard Revision.

TABLE II.
BINET-SIMON TESTS

Percentile Tables — Goddard 1911 Revision

3000 Cases

Percentiles	Year 5 180 Cases	Year 6 293 Cases	Year 7 370 Cases	Year 8 408 Cases	Year 9 425 Cases	Year 10 426 Cases	Year 11 338 Cases	Year 12 296 Cases	Year 13 184 Cases	Year 14 80 Cases
0	3.0	4.0	4.8	4.6	5.6	5.6	6.8	7.2	7.6	7.4
10	4.4	5.2	6.4	6.8	7.6	8.0	8.6	9.0	9.2	9.8
20	4.8	5.6	7.0	7.4	8.0	8.8	9.4	10.0	10.2	10.2
30	5.0	6.0	7.2	7.6	8.4	9.4	10.0	10.4	10.6	10.6
40	5.4	6.4	7.4	7.8	8.8	10.0	10.4	10.8	10.8	11.0
50	5.6	6.6	7.4	8.0	9.2	10.2	10.6	11.0	11.0	11.4
60	6.0	6.8	7.8	8.2	9.6	10.4	11.0	11.4	11.4	11.6
70	6.2	7.0	8.0	8.6	10.0	10.6	11.2	11.6	11.6	11.8
80	6.6	7.2	8.2	9.0	10.2	10.8	11.4	12.0	12.0	12.0
90	7.0	7.6	8.8	9.8	10.8	11.4	12.0	12.2	12.2	12.2
100	8.4	9.8	11.4	12.0	12.0	13.0	13.0	12.8	13.0	13.0

TABLE IV
STANFORD-BINET TESTS

Percentile Tables — Porteus 1916

945 Cases

Percentile	6-yr. chil. 71 cases	7-yr. chil. 124 cases	8-yr. chil. 152 cases	9-yr. chil. 140 cases	10-yr. chil. 119 cases	11-yr. chil. 114 cases	12-yr. chil. 127 cases	13-yr. chil. 98 cases
0	5^4	4^6	6^6	6^6	7^4	7^8	8^0	9^2
10	6^0	6^{10}	7^4	7^9	8^9	9^1	9^8	10^3
20	6^5	7^2	7^{10}	8^4	9^1	9^8	10^2	11^4
30	6^{10}	7^6	8^2	8^8	9^4	9^{10}	10^6	11^8
40	7^0	7^8	8^3	8^{10}	9^6	10^2	11^2	12^2
50	7^4	7^{10}	8^6	9^0	9^8	10^9	11^5	12^6
60	7^6	8^0	8^8	9^4	10^0	11^2	12^0	12^{10}
70	7^{10}	8^2	9^0	9^6	10^4	11^8	12^4	13^1
80	7^{11}	8^4	9^3	9^{10}	11^0	12^2	12^9	13^6
90	8^2	8^7	9^9	9^8	11^6	12^8	13^1	14^3
100	9^2	9^2	11^4	12^3	13^6	13^{10}	15^0	15^4

A consideration of the medians (fifty percentiles) will show that the medians of the mental ages of the five, six and seven year children corresponded well with the average chronological ages. On the other hand, the tests at the upper end of the Scale are so difficult that eleven year children have a median mental age nearly a year below their chronological age, whilst twelve and thirteen year tests are respectively eighteen months and over two years too difficult. For fourteen year children the difference between the actual age and the median test age is even greater, but this is partly due to the fact that the children were not given tests beyond the thirteen year level. Consequently, the average results were lower because some children would undoubtedly have earned credits from the tests above that level. The results, too, are undoubtedly affected by the fact that a large number of children of foreign-born parentage were included in this group. Probably these percentiles represent strata of the population more nearly comparable with the social grades from which the great body of defectives and delinquents is drawn.

An interesting point with regard to the distribution of ability is that when boys' and girls' results are separated and arranged in percentiles, the medians are practically identical. At eleven years the girls have an advantage of only .2 of a year. At year twelve they have a similar advantage, and at thirteen years an advantage of .4 of a year. Nor as regards range in ability is any appreciable difference noted. The girls range higher than the boys in four age levels, whilst the boys also range higher in four. These advantages are evidently due to chance. At the ten percentile point the boys have a higher test age in three years, as against the girls' advantage in three. Generally speaking, the tables show no significant sex difference.

For the benefit of those who have never used percentiles an approximate* method of arranging them may be described as follows: Supposing one thousand men were arranged in line, according to height, from the tallest at one end to the shortest at the other. The height of the shortest man in the whole array would be measured and put down as the zero percentile. The tallest man's measurement would equal the one-hundred percentile. The whole line would then be divided into ten equal groups of one hundred each. The one-hundredth man's height, counting from the shortest, would be the ten percentile, the two-hundredth, the twenty percentile, etc. The height of the middlemost man (the five-hundredth), would represent the fifty percentile, and this would constitute the median for the whole group. If the differences in the heights came down so evenly that a line joining the tops of their heads would be quite straight, then the median would exactly correspond to the average. An individual's percentile may be compared just as easily for mental measurements as for any other. When we wish to find what percentile the

*For exact method of arranging percentiles, see Yule's "Theory of Statistics," p. 150 *et seq.*

mental age of a given child equals, all we need do is to see what percentile his age corresponds to, using, of course, the table for his own chronological age group. For example, a twelve year child testing ten years of age mentally, per Binet, would be on the twenty percentile by Table II. That means that twenty per cent. of twelve year children are below this level. If we were 11.5 years mentally, then that age comes midway between the sixty and seventy percentiles. That means that he has a mental age greater than sixty-five per cent. of children at twelve years. For the purposes of diagnosis any child with a mental measurement coming below the ten percentile should be given further examination. The great advantage of a percentile table is that it not only enables us to see how far a child deviates from the median or average, but it also enables us to understand what significance should be attached to the amount of the deviation.

As regards the use of the Goddard Revision, the percentiles of Table II show that alterations in the standardization of the tests were very necessary. Medians for some years make it quite evident that the tests were too difficult; hence the publication of the Stanford Revision, with its easier standardization of the tests for the upper years, was particularly welcome. Nevertheless, results which have been obtained by the Goddard Revision, up to the level of twelve years, are quite reliable provided they are scored on a percentile basis in comparison with our table and not by mental age. With the provision of percentile tables for the Stanford Revision—a work upon which we are now engaged—an accurate comparison of Stanford and Goddard scores will be made possible. In this way a large body of discarded data will be available for research.

Modifications of the Scale

The alterations in position of the individual tests which we have deemed advisable to make are in consequence of the fact already mentioned, that the Stanford Revision was based on results from a socially selected group. Terman states that his subjects were drawn from schools in middle class areas in American cities. The chief use of the Binet lies in its application to subnormal and delinquent children. These are drawn, in the main, from a low social grade; hence, to apply a scale standardized for higher grade subjects may set up for defectives a somewhat different measure of average development. As previously explained, the results of the investigation reported in Table 1 were obtained with children of a lower social grade than Terman's. By this table it will be seen that the tests are somewhat too easy for the lower years, and too difficult for the upper years of the Scale. Accordingly, we have made such alterations n the standardization as our experience appeared most to warrant. We have

omitted other alterations that appeared desirable, as the provision of percentile tables will render many changes in standardization unnecessary.

The usual method of proving the correctness of standardization of individual tests has been to consider a test correctly placed in a year at which a certain percentage of children at that age pass the test. This percentage is usually 75, the idea being that if the test is correctly placed it will be passed by the 50 per cent. of children who may be considered of average intelligence, plus the 25 per cent. of children who are above average. Theoretically, it will not be passed by the 25 per cent. of children who are below average. It need hardly be pointed out that this is a plan based on purely theoretical considerations. Children who are rated by the whole Scale as being of the same mental age may have a very different development, as measured by their successes in the individual tests of the Scale. Hence it is quite possible that the 25 per cent. of children who fail in any single test may be made up of average and above average children as well as the dull. In the next place, it is very hard to obtain for standardization purposes a group of children in such well-balanced proportions of dull and bright children. The third consideration affecting the validity of the 75 per cent. standard is the very irregular development of mental functions with regard to chronological age. Ability, in certain directions, appears to develop quickly and regularly from one year of chronological age to another. Other abilities do not show such a marked improvement, one year being far too small a chronological unit of comparison. An examination of our table will show that the tests may be divided into two classes, the rapidly developing abilities, and those of slow or irregular development. It is noteworthy that the tests of memory are those in which the slow or irregular development is most apparent, whilst the tests of language and those which depend in the main on educational environment are those which show the most rapid improvement from year to year.

It does not follow, even if all the individual tests measure exactly up to the 75 per cent. standard, that the scale as a whole would be correct. Children do not earn their final score by successes in tests of the year of their mental age only, but are penalized by failures below their mental age, and gain advanced credits by the successes above it. Hence it is necessary to consider the percentage of passes achieved by children in any individual test over at least a three years' range in mental age. Table III gives the percentage of failures in each test by the three groups of children, the groups mentally a year above, a year below, and at the age in which the test is placed. For example, for each ten-year test the percentages of failures of children mentally nine, ten and eleven years are given, and these are then summarized in a separate column. These results were obtained with defectives and border-line cases, and from a comparison of the total failures we may judge the relative difficulty of the individual tests of the Stanford Binet series. By means of this data we have been enabled to

select the smallest number of tests in which alterations will have the greatest effect on the scale as a whole. The tests which defectives find most difficult are as follows:

V—2	Colors	123%	failures
V—5	Patience (Divided Rectangle)	150%	"
VI—6	Repetition of Sentences	156%	"
VII—3	Repeating 5 Digits	138%	"
VIII—1	Ball and Field	102%	"
VIII—4	Similarities	125%	"
VIII—6	Vocabulary	112%	"
IX—2	Weights	128%	"
IX—4	Repeating 4 Digits Backwards	180%	"
IX—6	Rhymes	100%	"
X—3	Memory for Designs	156%	"
X—6	Naming 60 Words	128%	"
XII—1	Vocabulary 40 Words	105%	"
XII—3	Ball and Field	120%	"

The alterations that we have made in the Scale on the basis of this classification are as follows: Tests IX—4 and X—3 are manifestly the most difficult of the series, and the former has, therefore, been advanced to Year X and the latter to Year XII; IX—Alt. 2 (Stamps) takes the place of IX—4. Test XII—3 shows an uneven development of ability, 50 per cent. of eleven-year chidren, 32 per cent. of twelve-year and no less than 38 per cent. of thirteen-year children fail in it. This is probably not so much due to the inherent difficulty of the test, but rather to the difficulty of scoring it correctly. We have, therefore, dropped it to the position of an alternative test for twelve years. XII—6 (repetition of 5 digits backwards) has also an irregular development of successes and is also placed as an alternative test. Test XII—5 (interpretation of fables) is also omitted because it shows a similar irregularity, but should, in our opinion, be retained as a sixteen-year test. These omissions and transpositions reduce the number of twelve-year tests to six, so that four months' credit should be allowed for each success. Since the omitted tests are the most difficult of the twelve-year series, these changes should have the desired effect of lessening the difficulty for this year. We have also made some alterations in the vocabulary tests, with the effect of making them easier for years ten, twelve and fourteen. We have made a selection of one-half the words which appear in Terman's full list. This selection has been made on the basis of an examination of a large series of vocabulary tests, in order to determine the relative difficulty of each word. It is noteworthy that our arrangement in order of difficulty did not agree with that given by Terman, and it also appeared that Terman's division of his full list into two series of supposedly equal difficulty did not agree with the experience obtained with defectives. Our results made it appear that Terman's first half

list was decidedly more difficult than his second half. Rating our subjects by our half list and doubling the score gave us approximately the same results as the actual scores on the full list, the only difference being that the children scored, almost uniformly, slightly better on the half list than on the full. Hence the test has been made somewhat easier, particularly for the years ten and twelve. The number of correct definitions required for success has been slightly modified. These alterations have been made on the basis of the examination of a large series of vocabulary tests, recording the average number of words correctly defined by the children at each mental age and counting the total number of times each word in the series was incorrectly defined.

At the lower end of the Scale we have made several alterations with the view to making the tests slightly more difficult. Test VII—4 (tying the bow-knot) depends more on chronological age than actual intelligence. Old, well-trained defectives score on the average a far greater percentage of successes than young children mentally advanced. It will be seen by Table III that defectives in the three mental age groups of six, seven and eight years only scored 19 per cent. of failures, as against a theoretical 100 per cent. Our results show that six-year children fail in 50 per cent. of the seven-year tests, seven-year children in 33 per cent. and eight-year children in 16 per cent.—equalling a total of 100 per cent. We have, therefore, put this test down as an alternative displacing it by Terman's alternative test 2 (repetition of 3 digits backwards). In Test VI—3 we allow only one trial for the counting of thirteen pennies; since the injunctions, "Put your finger on each one," is given in the instructions one trial seems sufficient. If one trial only is given, the test admittedly becomes a test of power of sustaining attention as well as of counting ability. It is not, therefore, so wholly dependent on previous training. In Test VI—5 the conditions of success are that four coins should be named correctly, whereas Terman allows three of four. The advanced chronological age of defectives evidently makes this test in the Stanford form too easy. These alterations to the lower tests of the Scale should have the effect of lessening the number of advance credits obtained by five-year children, and will also tend to diminish the effect of the previous training and thus the Scale will give a better measure of native ability.

A Condensed Diagnostic Scale

Doll, in arranging his brief scale, adopted a plan for determining the relative difficulty of single tests. He calculated the total number of successes in each year from five to ten years mental age, irrespective of the position of the test under examination.* No percentages of successes or failures by children above the mental age of ten years were given, a method which neglects the important fact that children mentally eleven and twelve years score a certain percentage

*See Doll, A Brief Binet-Simon Scale, Table I.

of failures in the tests of Year X. Unless a proper range of mental age is taken, one may be easily lead, as Doll is, into an unjustified conclusion that some of the ten and eleven, and all of the twelve-year tests are more difficult for normals than for defectives. (!) Doll's table proves that in these tests adult defectives scatter more than normal children, whether mentally average or advanced; and by comparing only one-half of the distribution it would naturally appear that defectives score better than normals. Doll remarks that to repeat six digits (Goddard X—3) is relatively easy for defectives, and to repeat seven digits (XII—1) proves to be very easy. In XII—1 the percentage of successes by normal subjects mentally five, six, seven and eight years is zero, a result wholly to be expected, since children of these chronological ages are not likely to pass a twelve-year test. On the other hand, it is not at all unusual for adult defectives, mentally eight, nine and ten, to score a proportion of twelve-year successes. Had the comparison continued beyond the mental age of ten years, the result would have stood differently. It has been necessary to point this out in order to show that our list giving the tests which are difficult for defectives has not been arranged by the same method as Doll has adopted in giving his table showing what he calls "Differential Difficulty for Feeble-minded Subjects." It will be noted that we have not instituted any comparison between adult defectives and normal children. In our opinion, Table III is sufficient to show the relative difficulty of the tests for defectives.

From our list has been selected a series of eight tests which are amongst those most commonly failed by defectives. As our table is limited to a three years' mental range, our method excludes from calculations failures made by atypical cases who show a very irregular development. By comparing the tests of a doubtful or border-line case with the typical failures of the definitely defective, we may increase the diagnostic value of the Scale. In other words, our list is not a brief scale to permit of a quick and fairly accurate estimate of the mental age, but is an examination of the subject's results in order to discover from what individual tests he obtained his mental age score. If he shows the characteristic failures of defectives, the presumption of feeble-mindedness is thereby increased. The rule that we have adopted is to count the diagnostic score by allowing a year for each success achieved by the subject in the eight tests of the scale which we have selected on the basis of difficulty for defectives, beginning to count from six years. For example, one success gives six years, two successes seven years in diagnostic score, and so on. The tests are as follows: V—5 (Patience), VII—3 (Repeating 5 Digits), VIII—1 (Ball and Field), IX—2 (Weights), X—3 (Repeating 4 Digits Backwards), X—6 (60 Words), XII—3 (Memory for Designs), XII—1 (Vocabulary), or alternatively XII—Alt. 2 (Ball and Field). The procedure is to give both the alternative tests for year twelve, but only to count one success even when both are passed.

TABLE III.

Percentage of Failures in Individual Tests over a Three Years' Mental Range.

Test	Mental Age	No. 1	2	3	4	5	6	
Year V	4 years	51	74	65	20	62	41	
	5 "	18	31	27	9	63	22	
	6 "	9	18	6	3	25	22	
Total Percentage Failures		78%	123%	98%	32%	150%	85%	
Year VI	5 years	64	50	50	45	50	86	
	6 "	18	23	7	29	9	56	
	7 "	1	1	3	11	3	14	
Total Percentage Failures		83%	74%	60%	85%	62%	156%	
Year VII	6 years	31	37	70	9	42	54	
	7 "	10	40	47	7	7	17	
	8 "	3	3	21	3	9	9	
Total Percentage Failures		44%	80%	138%	19%	58%	80%	
Year VIII	7 years	56	56	31	69	54	74	
	8 "	31	25	22	47	22	22	
	9 "	15	9	18	9	9	6	
Total Percentage Failures		102%	80%	71%	125%	85%	112%	
Year IX	8 years	25	47	42	77	37	58	
	9 "	18	39	18	53	9	33	
	10 "	9	42	10	50	5	9	
Total Percentage Failures		52%	128%	70%	180%	51%	100%	
Year X	9 years	53	43	66	26	33	60	
	10 "	25	40	45	25	10	45	
	11 "	18	4	45	9	0	23	
Total Percentage Failures		96%	87%	156%	60%	43%	128%	

		No. 1	2	3	4	5	6	7	8
Year XII	11 yrs.	55	41	50	23	27	36	18	41
	12 "	42	21	32	5	17	21	32	26
	13 "	8	8	38	0	31	38	31	8
Total Percentage Failures		105%	70%	120%	28%	75%	95%	81%	75%

It will be noted that these eight tests cover a very wide range of capacities and that all of them except XII—1 (Vocabulary) are unaffected by previous educational environment. In other words, they represent tests of native capacity rather than of acquirement. The assumption appears to be warranted that children who score a large proportion of successes in these selected tests are those with the best native ability, and if socially unfit, are so on account of deficiencies other than intellectual.

In order to see how far such an assumption is justified a group of seventy of the highest grade cases at The Training School were rated according to their degree of feeble-mindedness as determined by their institution history. These ratings were given by Mrs. Nash, the school principal, and were independent of the mental-age scores.

The Pearson correlation between these children's "diagnostic scores" in the selected tests and Mrs. Nash's ratings was .70. Considering that the group represented the most difficult cases for diagnosis and contained dull normals, border-line cases, and high-grade defectives, this correlation must be considered high. Had a wider range of mental ability been taken, the correlation would have been sensibly increased.

Of those rated by Mrs. Nash as being dull normals 29, or 94 per cent., scored above 10 years in "diagnostic score,' or if less than 13 years in chronological age made less than two failures in the selected list. Of those rated as border-line cases, or of doubtful diagnosis, 10, or 83 per cent., also satisfied these conditions of success. Of those considered definitely feeble-minded only seven, or 29 per cent., succeeded. Hardly one of these seven was a typical feeble-minded case. One was a paralytic with a remarkable though irregular mental development. Two show symptoms decidedly psychopathic. Three are of high mentality, but are morally degenerate, and a seventh case is mentally well developed, but is unemployable entirely on account of temperamental peculiarities.

The conclusion that these results appear to warrant is that the "diagnostic scale" serves to indicate the cases who have attained their mental age rating by reason of good native ability. It is also a material aid to diagnosis. A child, apparently on the border-line of mental deficiency, would have the presumption of his normality decidedly increased by rating 11 years or more in "diagnostic score."

On the other hand, the results of extended expert observation prove more than ever that, used alone, the Binet can never be a satisfactory instrument for diagnostic purposes. There are too many "negative instances" wherein the person's social efficiency or the reverse has been shown to be in no wise dependent on Binet age. The "diagnostic score" suffers the same handicap. All that is claimed for it is that it gives added meaning to Binet age in a large percentage of cases in which diagnosis is most difficult.

Part II

Tests and Test Procedure

By S. D. Porteus and Helen F. Hill

In this booklet we have endeavored to present the Binet-Simon tests in a more usable form than they have been hitherto available. The plan has been to set out the tests concisely, and to give the exact procedure and conditions of success in as handy a form for reference as possible. In order to further the last consideration, the tests themselves have been freed of all discussion, either as regards the standardization or the psychological principles involved. Such discussion is to be found in full in Terman's book, "The Measurement of Intelligence," to which we would refer the reader. Points upon which we are not in agreement with Terman, or alterations in the form of the tests, have been discussed by us in a distinct section under the heading, "Notes and Comments."

Terman's book is more of a text book for study than a reference guide to the application of the tests themselves. On the other hand, the "Record Blank" arranged by Terman, though it provides a key to the scoring and application of the tests, is insufficient, as the exact procedure for each individual test is not given. Either the examiner must trust to his memory or have Terman's book at his elbow the whole time. Even with experienced examiners it is unwise to trust entirely to memory. In Terman's bock the tests themselves are scattered throughout two hundred pages of discussion and explanation. Regarding, as we do, "The Measurement of Intelligence" as a textbook rather than a reference handbook, we feel that there is a need for a more concise form of the tests. This guide is not intended to supersede the use of the book, which should be read and reread by everyone who applies the tests.

As regards modifications which we have introduced, they will be found to consist mainly of alterations in phraseology, bringing the language of the tests more within the child's range of comprehension. As such we believe them to be improvements in procedure, but in most cases we have been careful not to make alterations which would affect the standardization of the tests. In other words, whilst our alterations are intended to make certain tests more intelligible to the child's comprehension, they do not either lessen or increase their difficulty. A number of alterations in the procedure which we thought desirable have not been made, because they might have affected the standardization of the tests. Tests restandardized have already been recorded in the previous section.

TESTS IN NEW ORDER

Year III.
1. Points to parts of body.
2. Names familiar objects.
3. Enumerating objects, pictures.
4. Giving sex.
5. Giving family name.
6. Memory for sentences.

Alt. Repeating 3 digits.

Year IV.
1. Comparison of lines.
2. Discrimination of forms.
3. Counting 4 pennies.
4. Copying square.
5. Comprehension.
6. Repeating 4 digits.

Alt. Memory for sentences.

Year V.
1. Comparison of weights.
2. Naming colors.
3. Aesthetic comparison.
4. Simple definitions.
5. Divided rectangle.
6. Memory for commissions.

Alt. Giving age.

Year VI.
1. Knowing right and left.
2. Missing features.
3. Counting 13 pennies.
4. Question comprehension.
5. Naming coins.
6. Memory for sentences.

Alt. Morning or afternoon.

Year VII.
1. Number of fingers.
2. Picture description.
3. Repeating 5 digits forwards.
4. Repeating 3 digits backwards.
5. Differences.
6. Copying diamond.

Alt.
1. Days of week.
2. Tying a bow-knot.

Year VIII.
1. Ball and field.
2. Counting backwards.
3. Question comprehension.
4. Similarities.
5. Definitions.
6. Vocabulary.

Alt.
1. Naming 6 coins.
2. Writing from dic...

Year IX.
1. Knowing the date.
2. Arranging weight.
3. Making change.
4. Counting stamps.
5. Sentence compos...
6. Making rhymes.

Alt. Naming months.

Year X.
1. Vocabulary.
2. Detecting absurdities.
3. Repeating 4 digits backwards.
4. Reading and memory.
5. Question comprehension.
6. Naming 60 words.

Alt.
1. Repeating 6 digits.
2. Memory for sentences.

Year XII.
1. Vocabulary.
2. Defining abstract words.
3. Memory for designs.
4. Dissected sentences.
5. Picture interpretation.
6. Similarities.

Alt.
1. Repeating 5 digits backwards.
2. Ball and field.

Year XIV.
1. Vocabulary.
2. Repeating 7 digits.
3. Differences—king and president.
4. Problems of fact.
5. Arithmetical reasoning.
6. Reversing of clock hands.

Alt. Paper folding.

Year XVI.
1. Vocabulary.
2. Interpretation of fables.
3. Defining abstract terms.
4. Enclosed boxes.
5. Repeating 6 digits backwards.
6. Code test.

Alt.
1. Memory for sentences.
2. Physical relations.

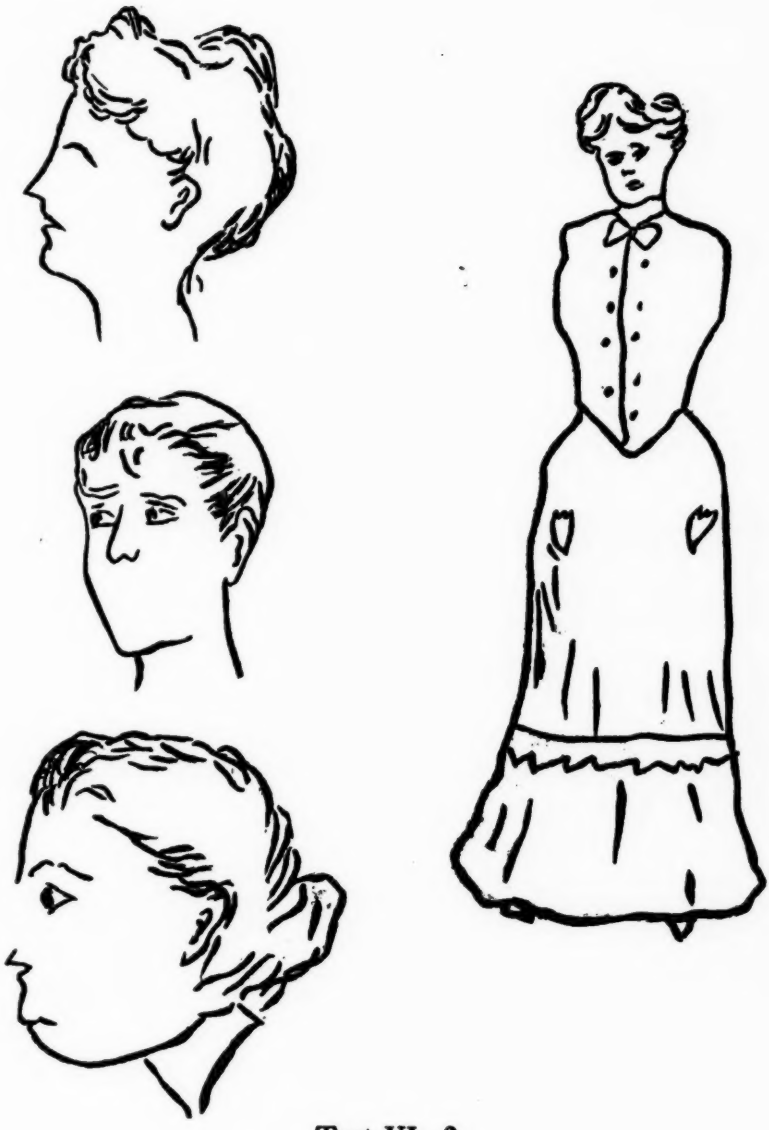
TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
III 1 Pointing to parts of the body.	*"Show me your nose. Put your finger on your nose."* Repeat command for eyes, mouth, and hair. If child is unresponsive, or shy, examiner may point to other features and say, *"Is this your nose?"*	Child points to three of four correctly.
2 Naming familiar objects: large door key, penny, knife, lead-pencil.	*"Tell me what this is."* Examiner shows: (a) *Key,* (b) *Penny,* (c) *Closed Knife,* (d) *Lead-Pencil.*	Child names three out of five objects.
3 Enumerating objects: Pictures as listed or alternatively —Shoemaker, Janitor, Butcher, as used in Goddard Revision.	*"Tell me what you see in this picture."* Shows: (a) *Dutch Home.* (b) *Canoe.* (c) *Post Office.* If no response, then say, *"Look at the picture and tell me everything you can see in it."* If still no response, say, *"Show me the ———,"* (naming some object). Only one question of this type is allowed. If child stops before giving three objects he may be urged by saying, *"And what else?"* (See note 1.)	Child names at least three objects in one picture.
4 Giving sex.	(For boys), *"Are you a little boy or a little girl?"* (For girls), *"Are you a little girl or a little boy?"*	Correct answer.
5 Giving family name.	*"What is your name?"* If first name given, for example, *"Walter,"* then say, *"What is your other name? Walter—what?"* If child fails to respond examiner may suggest fictitious name.	Correct answer.
6 Memory for sentences.	*"Can you say—'Nice Kitty?'"* *"Now say after me:* (a) *"I have a little dog."* (b) *"The dog runs after the cat."* (c) *"In summer the sun is hot."* If child fails to respond at all, the first sentence only may be repeated two or three times. The first sentence in this case counts as a failure.	One sentence without error after single reading.
Alt. Repeating three digits.	*"Say after me:* (a) *6-4-1,* (b) *3-5-2,* (c) *8-3-7."* One second interval in reading.	One set correct in right order after single reading.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
IV 1 Comparison of lines: Card with lines differing one centimetre in length.	*"Look at these two lines. Now put your finger on the long one."* (b) After turning card: *"Now show me the long one."* (c) Turn card again: *"Now show me the long one."* (See note 2.)	Three correct responses out of three trials. (Position of card altered each trial.)
2 Discrimination of forms: Card with forms.	See Terman's "Measurement of Intelligence," or use alternative test below.	Seven of ten forms matched correctly.
3 Counting four pennies.	*"Count these pennies, and tell me how many there are. Put your finger on each one as you count them, like this—One."*	Correct response. Insure counting aloud, and touching with fingers at same time.
4 Copying square: Card with square. (See Record Blank.)	*"You see this"* (pointing to square), *"I want you to make one with the pencil just like it. Draw it here."* Give three trials, repeating directions and pointing out copy each trial.	One of three trials recognizable as a square. Right angles more important than closed corners or straight lines.
5 Comprehension.	*"What must you do?"* (a) *"When you are sleepy?"* (b) *"When you are cold?"* (c) *"When you are hungry?"* Questions may be repeated if child fails to answer. (See note 3.)	Satisfactory answers to two out of three questions. (Terman allows about 20 seconds for each answer.)
6 Repeating four digits.	*"I am going to say some numbers, and after I have done, I want you to say them just as I have said them. Listen:* (a) *"4-7-3-9,"* (b) *"2-8-5-4,"* (c) *"7-2-6-1."* If child begins to repeat before all four digits are given say, *"Don't say anything till I'm done,"* and begin a new set. (See note 4.)	One of three correct. (Read one per second.)
Alt. Memory for sentences.	*"Say after me:"* (a) *"The boy's name is John. He is a very good boy."* (b) *"When the train passes you will hear the whistle blow."* (c) *"We are going to have a good time in the country."*	One sentence without error after single reading, or two sentences with one error each.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
V 1 Comparison of weights: Two blocks of equal size, 3 and 15 grams weight.	*"You see these two blocks? Now, one of them is heavy and the other is light. Lift each one up, and tell me which is the heavy one."* If test is not understood, examiner may illustrate process of trying weight of each block by lifting from the table.	Two out of three trials correct, blocks being moved into alternate positions for each trial. See that child compares blocks for each trial.
2 Naming colors. Card with colors, red, yellow, blue, green.	*"I want you to tell me the names of these colors. What color is that?"* (See note 5.)	No error in naming.
3 Aesthetic comparison: Pairs of pretty and ugly pictures. See Figure I.	Show pictures in pairs. *"Look at these two pictures. Now, which is the pretty one?"* Repeat question for each pair. (See note 6.)	No error.
4 Simple definitions.	*"You have seen a chair. You know what a chair is. Tell me, what is a chair?"* Repeat question for: (b) *Horse,* (c) *Fork,* (d) *Doll,* (e) *Pencil,* (f) *Table.*	Child defines in terms of use, or describes the object. Four out of six correct.
5 Divided rectangle: Two similar oblong cards, one uncut, the other cut diagonally.	Place cards on table, thus: *"I had a card and cut it into two pieces. Here they are."* (Indicate cut pieces.) *"I want you to take these two pieces and put them together so that they will look just like this"* (pointing to uncut card). Repeat instructions if first attempt fails.	Two successes in three trials. Child to be prevented from turning one piece over.
6 Memory for commissions: Key, chair, and box. (Care should be taken that objects are in different parts of the room.)	*"I want you to do something for me. Do you see this key? I want you to put it on that chair, then open the door, and then bring me the box that you see over there. Now, first the key on the chair, then open the door, then bring me the box."* Emphasize *"first"* and *"then."*	Three commissions correct and in right order; no further help or prompting than in instructions.
Alt. Giving age.	*"How old are you?"*	Correct response.

Test V—3

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
VI 1 Knowing right and left.	*"Show me your right hand. Show me your left ear. Show me your right eye."* If one error repeat test, using *left hand, right ear, left eye.*	Three of three correct in first test, or five out of six, if six questions given.
2 Missing features. Pictures. See Figure 2.	Show first picture— *"There is something wrong with this face; something has been left out. Look carefully and tell me what part of the face is not there. This one has no ———?"* If child says that some other part of the body is missing, say—*"No, look at the face."* (b and c) *"What is left out of this face?"* (d) *"What is left out of this picture."* (See note 7.)	Three of four correct. If child fails on first test it is allowable to point out in this picture the missing feature, before proceeding to the other questions.
3 Counting thirteen pennies.	Proceed as in Test IV—3.	Correct answer—one trial.
4 Question comprehension.	(a) *"What ought you to do: "If it is raining when you start for school?"* (b) *"What ought you to do: "If you find that your house is on fire?"* (c) *"What ought you to do: "If you are going somewhere and miss your train (or car)."*	Two out of three correct. Question may be repeated if child does not understand.
5 Naming coins: nickel, penny, quarter, dime.	Show *nickel, penny, quarter, dime* in given order. Say for each: *"What is that?"* If child says *"money,"* say: *"But what do you call this piece of money?"*	All correct.
6 Memory for sentences.	*"Listen, I am going to say some sentences and when I have finished I want you to say just exactly what I have said:* (a) *"We are having a fine time. We found a little mouse in the trap."* (b) *"Walter had a fine time on his vacation. He went fishing every day."* (c) *"We will go out for a long walk. Please give me my pretty straw hat."*	One sentence repeated without error, or two with not more than one error each. Sentence to be given once only.
Alt. Morning or afternoon.	If morning—*"Which is it, morning or afternoon?"* If afternoon—*"Which is it, afternoon or morning?"*	Correct response.



Test VI—2

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
VII 1 Number of fingers.	*"How many fingers have you on your right hand? How many on your left? How many on both hands together?"* Credit if number of fingers given excluding thumbs.	Correct response—child to be prevented from counting.
2 Picture description: See pictures used in Test III-3.	Show *Dutch Home.* *"Tell me all about this picture. What is this a picture of?"* (b) Repeat instructions for *Canoe.* (c) Repeat instructions for *Post Office.* (See note 8.)	Two out of three pictures described in terms of action rather than of enumeration.
3 Repeating five digits.	See directions Test IV—6. *"Say after me,* (a) *3-1-7-5-9,* (b) *4-2-8-3-5,* (c) *9-8-1-7-6."* Read about one digit per second.	One of three correct. Each set to be read once only.
4 Repeating three digits backwards.	*"I am going to say some numbers to you, but this time I want you to say them backwards. If I should say 6-2-5, you would say 5-2-6. Ready? Now listen carefully:* (a) *2-8-3,* (b) *4-2-7,* (c) *9-5-8."*	One of three correct.
5 Differences.	*"What is the difference between a fly and a butterfly?"* (If no response)—*"You have seen a fly and a butterfly, haven't you? Now, tell me in what way a butterfly is not like a fly."* (b) Repeat question for: *Wood and glass.* (c) Repeat question for: *Stone and egg.* (See note 9.)	Two of three differences given. Common differences accepted.
6 Copying diamond. Card with diamond. Pen. See Record Blank.	Show child the model drawing. Say: *"Look at this. I want you to draw one exactly like it."*	Two of three trials correct. Easily recognizable shape.
Alt. Days of week.	*"You know the days of the week. Name all the days of the week for me."* Checks: *"What day comes before Tuesday? What day comes before Thursday? What day comes before Saturday?"*	Correct order for days. Two of three checks correct.
Alt. 2 Tying a bow-knot. Bow-knot tied on stick.	Show string tied in plain bow-knot around stick. Say: *"Here is a bow-knot. I want you to take this other piece of string and tie the same kind of knot around this pencil."* (See note 10.)	Correct knot. Half credit for single bow. (Terman allows one minute.)

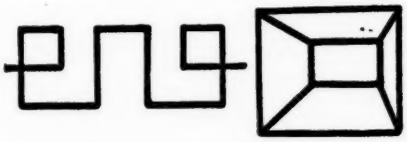
TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
VIII 1 Ball and field. See Record Blank.	*"Now suppose that a ball has been lost in this big round field. We don't know where it is or anything about it, except that it is lost somewhere in the field. Now, suppose you had to go and find it. I want you to take this pencil and show me how you would go about looking for the ball so as to be sure not to miss it. Begin at the gate and mark for me the path you would take."* If child stops at any given point before covering the ground then say, *"But suppose you have not found it yet. Now where would you go?"* (See note 11.)	Test is passed if subject shows evidence of pursuing any definite plan. (Insure that he does not lift his pencil from paper until test is completed.)
2 Counting backwards.	*"Do you know how to count backwards? I want you to count backwards for me from 20 to 1."* (See note 12.)	Child counts with not more than one error of omission or transposition. (Terman allows 40 seconds. Errors self-corrected do not count.)
3 Question comprehension.	(a) *"What ought you to do:* *"When you've broken something which belongs to someone else?"* (b) *"What ought you to do:* *"When you're on your way to school and notice that you're in danger of being late (tardy)?"* (c) *"What ought you to do:* *"If a boy (girl) hits you without meaning to do it?"* Question may be repeated.	Two of three satisfactory answers.
4 Similarities.	*"I am going to name two things which are like one another in some way, and I want you to tell me how they are alike."* (a) *"Wood and coal. In what way are they alike?"* (b) *"In what way are an apple and a peach alike?"*	Likeness given in two of four comparisons.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
	(c) *"In what way are iron and silver alike?"* (d) *"In what way are a ship and an automobile alike?"* If a difference is given say, *"No, I want you to tell me how they are alike. In what way are.....and..... alike?"*	
5 Definitions.	(a) *"What is a balloon?"* Repeat question for (b) *tiger*, (c) *football*, and (d) *soldier.* If child gives the class to which the object belongs without further description, as for example: *"A soldier is a person,"* or *"A soldier is a man,"* say—*"You have not told me enough about it,"* and repeat the question —*"What is a soldier?"* (See note 13.)	Two of four words defined in terms superior to use. Definitions which involve a description of the object, composition, or class to which it belongs, plus some "defining mark" are satisfactory.
6 Vocabulary: List of fifty words—see Record Blank.	*"I want to find out how many words you know the meaning of. When I say a word you tell me what it means:* *"What is an orange? Tell me any way you please."* If child uses the word correctly in a sentence or in any other way gives response of doubtful value, say: *"Explain what you mean."* (See note 14.)	Ten satisfactory definitions.
Alt. 1 Naming six coins: Nickel, penny, quarter, dime, dollar, half-dollar.	Show *nickel, penny, quarter, dime, dollar, half-dollar* in order. Say: *"What is that?"* If child says *"money,"* say, *"But what do you call the piece of money?"*	All six correctly named.
Alt. 2 Writing from dictation.	Give child pen, ink, and paper and say: *"I want you to write something for me as nicely as you can. Write these words, 'See the little boy.' Be sure to write it all, 'See the little boy.'"*	Passed if written legibly.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
IX 1 Knowing the date.	(a) *"What day of the week is it today?"* (b) *"What month is it?"* (c) *"What day of the month is it?"* (d) *"What year is it?"*	Correct answer for a, b, and d. An error of three days in either direction allowed in c. Self-correction allowed.
2 Arranging weights. Blocks of identical appearance, 3, 6, 9, 12, and 15 grams.	*"You see these little blocks. None of them weigh the same. One of them is the heaviest of all. I want you to find it and put it here. Then find the next heaviest and put it here, and then the next heaviest and then the next here, and the lightest here. Remember, the heaviest of all here, etc."* (Repeat instructions.) (See note 15.)	Weights correctly arranged. Two of three trials. Do not illustrate procedure of weighing to the child, but allow him to discover his own method.
3 Making change.	(a) *"If I were to buy four cents' worth of candy and gave the store-keeper ten cents, how much money should I get back?"* With same form of question (b) *15c-12c.* With same form of question (c) *25c-4c.* (See note 16.)	Two of three correct. (Terman allows 15 seconds for each problem.)
4 Counting stamps: Card with 3 one-cent and 3 two-cent stamps.	Show card with three one-cent and three two-cent stamps. *"Look, these are one-cent stamps, and these are two-cent stamps. How much money would you need to buy all these stamps?"*	Correct answer in one trial.
5 Sentence composition.	(a) *"You know what a sentence is. A sentence is like a little story that you tell about anything. Now, I am going to give you three words and you must make up a sentence or little story with all three words in it. The three words are 'boy,' 'ball,' 'river.' Remember, you must use all the words in the one sentence.* (b) *"Now make a sentence using 'work,' 'money,' 'men.'* (c) *"Now make a sentence using 'desert,' 'rivers,' 'lakes.'* (See note 17.)	Two of three sentences correct. Compound or complex sentences must not contain more than two distinct statements.
6 Making rhymes.	*"You know what a rhyme is. A rhyme is a word that sounds like another word. The two words 'hat' and 'rat' sound alike, and so they make a rhyme. Hat, rat, cat, bat, all rhyme with one another. Now give me all the words you can think of in one minute that rhyme with 'day.'"* (b) *"Now all the words that rhyme with 'mill.'"* (c) *"Now all the words that rhyme with 'spring.'"*	Three rhymes given to two of three words. If child does not succeed with "day," examiner may give rhymes before proceeding with "mill" and "spring."
Alt. Naming months.	*"Name all the months of the year."* Checks: *"What month comes before April?"* *"What month comes before July?"* *"What month comes before November?"*	Not more than one error in naming in order. Two of three checks correct. (Terman allows 20 seconds.)

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
X 1 Vocabulary. See Record Blank.	Proceed as in Test VIII—6.	Fourteen words defined correctly.
2 Detecting absurdities.	*"I am going to read a sentence which has something foolish in it, some nonsense. I want you to listen carefully and tell me what is foolish about it."* After reading—*"What is foolish about that?"* (a) *"A man said, "I know a road from my house to the city which is down hill all the way to the city and down hill all the way back home.'"* (b) *"An engineer said that the more cars he had on his train the faster he could go."* (c) *"Yesterday the police found the body of a girl cut into eighteen pieces. They believe that she killed herself."* (d) *"There was a railroad accident yesterday, but it was not very serious. Only forty-eight people were killed."* (e) *"A bicycle rider, being thrown from his bicycle in an accident, struck his head against a stone and was instantly killed. They picked him up and carried him to the hospital, and they do not think he will get well again."*	Four of five correct. If in doubt as to correctness of reply, child may be asked to explain his answer.
3 Repeating four digits backwards.	*"I am going to say some numbers to you, and when I have finished I want you to say them backwards. If I should say 5-1-9 you would say 9-1-5. Listen:* (a) *6-5-2-8,* (b) *4-9-3-7,* (c) *8-6-2-9."*	One of three correct. Read at rate of about one digit per second. One reading only.
4 Reading and memory. Printed selection. See Record Blank.	*"I want you to read this for me as well as you can."* After reading, take selection away and say: *"Very well done. Now I want you to tell me what you have read. Begin at the beginning and tell everything that you can remember."* (See Terman's note, page 262.) (See note 18.)	Not more than two errors in reading. Child must reproduce at least eight "memories." Exact reproduction not required so long as idea is given. (Terman allows 35 seconds for reading.)

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
5 Question comprehension.	(a) *"What ought you to say when someone asks you your opinion of a person you don't know very well?"* (b) *"What ought you to do before beginning something very important?"* (c) *"Why should we judge a person more by his actions than by his words?"*	Two of three satisfactory answers. Question may be repeated.
6 Naming sixty words.	*"I want to see how many words you can say for me in three minutes. When I say 'ready' you must begin and give all the words you can think of as fast as you can. Words such as clouds, dog, chair, happy—any words at all will do."*	Names sixty words in three minutes without counting repetitions. Half credit if over fifty-five named.
Alt. 1 Repeating six digits.	Proceed as in VII—3 (a) *3-7-4-8-5-9,* (b) *5-2-1-7-4-6.*	One of two, order correct.
2 Memory for sentences.	Proceed as in VI—6. (a) *"The apple tree makes a cool, pleasant shade on the ground where the children are playing."* (b) *"It is nearly half-past one o'clock; the house is very quiet and the cat has gone to sleep."* (c) *"In summer the days are very warm and fine; in winter it snows and I am cold."*	One of three correct, or two with one error each.

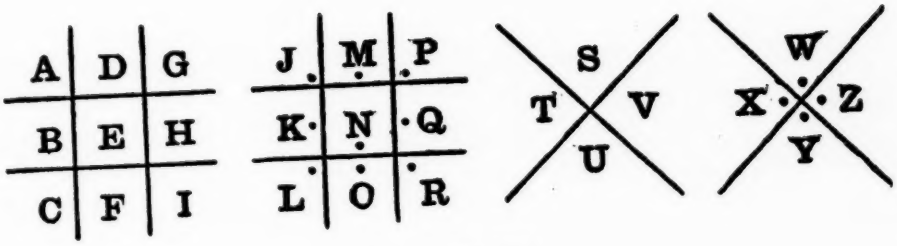


Test XII—3

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
XII 1 Vocabulary. See Record Blank.	Proceed as in Test VIII—6.	Eighteen definitions.
2 Defining abstract words.	(a) *"What is pity? What do we mean by pity?"* Repeat question for: (b) *Revenge,* (c) *Charity,* (d) *Envy,* (e) *Justice.*	Three of five words correctly defined. Definition by illustration accepted.
3 Memory for designs. See Fig. 3.	*"This card has two drawings on it. I am going to show them to you for ten seconds; then I will take the card away and let you draw from memory what you have seen. Look at both drawings carefully and remember that you have only ten seconds."*	One correct and the other with only one mistake such as omission or transposition of a line or the placing of the inner rectangle in "b" in the center instead of a little off center.
4 Dissected sentences. See Record Blank.	*"I had a sentence and I changed the words around so that they don't make sense. Here they are:* (Read words of first sentence.) *Now, I want you to look carefully at these words and see if you can tell me what the sentence was. Be sure to use all the words, but don't use any others."*	Two of three correct. If unsuccessful with first sentence in one minute, illustrate how words may be recombined to make sense. (Terman allows one minute for each.)
5 Picture interpretation. Use Terman's four pictures.	*"Tell me what this picture is about, explain this picture."* (See note 19.)	Three of four correctly interpreted.
6 Similarities.	*"In what way are a snake, a cow, and a sparrow alike? Tell me some thing which is true of all of them."* (b) *Book, teacher, newspaper.* (c) *Wool, cotton, leather.* (d) *Knife-blade, penny, piece of wire.* (e) *Rose, potato, tree.* (See note 20.)	Three of five correct — superficial resemblances accepted.
Alt. 1 Repeating five digits backwards.	Proceed as in Test VII—4 and Test X—3. (a) *3-1-8-7-9,* (b) *6-9-4-8-2,* (c) *5-2-9-6-1.*	One of three trials correct.
Alt. 2 Ball and field. See Record Blank.	Proceed as in VIII—1.	Superior plan. Spiral or transverse parallel lines joined at ends.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
XIV 1 Vocabulary. See Record Blank.	See Test VIII—6, Test X—1, Test XII—1.	Twenty-four words correctly defined.
2 Repeating seven digits.	Procedure as in year X—Alt. 1. (a) *2-1-8-3-4-3-9,* (b) *9-7-2-8-4-7-5.*	One of two correct.
3 Differences: king and president.	*"There are three main differences between a king and a president. What are they?"* (Terman accepts as answer either that the king is more powerful than a president or vice versa).	Two of three differences given. Only correct responses are those which refer to power, accession, tenure.
4 Problems of fact.	*"I am going to read you something. Listen very carefully:* (a) *'A man who was walking in the woods near a city stopped suddenly, very much frightened, and then ran to the nearest policeman, saying that he had just seen hanging from the limb of a tree—a what?'* (b) *'My neighbor has been having queer visitors. First a doctor came to his house, then a lawyer, then a minister (preacher or priest). What do you think happened there?'* (c) *'An Indian who had come to town for the first time in his life saw a white man riding along the street. As the white man rode by the Indian said—'The white man is lazy; he walks sitting down.' What was the white man riding on that caused the Indian to say 'he walks sitting down'?"* (See Terman, page 315.)	Two of three correct responses. To second, child is expected to state why each visitor came.
5 Arithmetical reasoning.	Child reads from record blank: (a) *"If a man's salary is $20.00 a week and he spends $14.00 a week, how long will it take him to save $300.00?"* (b) *"If two pencils cost five cents, how many pencils can you buy for fifty cents?"* (c) *"At fifteen cents a yard, how much will seven feet of cloth cost?"* (See Terman, page 319.)	One minute allowed for solution of each problem. Two of three correct.
6 Reversing hands of the clock.	*"Suppose it is twenty-two minutes after six o'clock. Can you see in your mind where the large hand would be, and where the small hand would be? Now, suppose the large hand and the small hand were to change places, so that the large hand is where the small one was, and the small one where the large one was. What time would it be then?"* Repeat test with hands at (b) *8.10,* and (c) *2.46.*	Two of three correct with following ranges of accuracy: (a) 4.30 to 4.35 (b) 1.40 to 1.45 (c) 9.10 to 9.15
Alt. Paper folding.	(See Terman, page 310.)	Test passed if rule is grasped by time sixth sheet is reached.

TEST DESCRIPTION AND MATERIAL	PROCEDURE	CONDITIONS OF SUCCESS
XVI 1 Vocabulary. See Record Blank.	See previous tests—VIII—6.	Thirty-two words.
2 Interpretation of Fables.	*"You know what a fable is. A fable is a little story which is meant to teach us a lesson. Now, I am going to read a fable to you and when I have finished I am going to ask you what lesson the fable teaches us."* (See Terman, page 290.) After reading—*"What lesson does that fable teach us?"*	Eight points, four correct or equivalent in half credits. Half credit if lesson is given in generalized form but not exactly correct interpretation.
3 Defining abstract terms.	*"What is the difference between:* (a) *Laziness and idleness?* (b) *Evolution and revolution?* (c) *Poverty and misery?* (d) *Character and reputation?"* (See Terman' page 324.)	Three of four correct. Answer must indicate a real contrast.
4 Enclosed boxes problem.	(See Terman, page 327.)	Three of four correct.
5 Repeating six digits backwards.	*"Say after me backwards:* (a) *4-7-1-9-5-2.* (b) *5-8-3-2-9-4.* (c) *7-5-2-6-3-8."* See Test X—3.	One of three correct. One reading.
6 Code Test. Form giving code.	*"These diagrams contain all the letters of the alphabet arranged in this way.* (Point out each letter.) *In these first two diagrams the letters go up and down and in the other two they run in the opposite way to the hands of the clock.* (Point out.) *You see that the second diagram is just like the first except that every space has a dot in it, and the last diagram is like the third except that here also each letter has a dot.* *"From these diagrams can be made a code, a secret language, which was used in the Civil War. To make the code we use the lines which hold the letter instead of the letter itself."* (Codify 'spy' and 'war.') *"Now, I want you to write me a message with the code."* Point out again position of letters and the dots. Take the diagram away and tell subject to write, *"Come quickly,"* in the code.	Words written in six minutes without more than two errors. Dot omitted counts half error. Subject must not be allowed to write whole code and then copy.
Alt. 1 Memory for sentences.	(a) *"Walter likes very much to go on visits to his grandmother, because she tells him many funny stories."* (b) *"Yesterday I saw a pretty little dog in the street. It had curly hair, short legs, and a long tail."*	One sentence correct without any error.
Alt. 2 Physical relations.	See Terman "The Measurement of Intelligence," page 333.	



Test XVI—6

Part III

Notes and Comments

(1) Year III—3: *Enumeration of Objects in Pictures.* In this test we recommend the use of three of the pictures used with the Goddard Revision. These have certain advantages over Terman's selections. They are colored and thus more likely to interest the young child. Secondly, pictures such as Terman's, representing situations with a not too obvious meaning and thus difficult enough to serve as an interpretation test for Year XII, are rather unlikely to be suitable for an enumeration of objects test in Year III. As a matter of fact the three pictures used by Terman are of very uneven value for the latter purpose. The "Dutch Home" contains more distinct objects that are familiar to a small child than either the "River Scene" or the "Post-office." Experiments to show the number of distinct objects that can be enumerated in ten seconds prove than the best selection would be the "Dutch Home," the "Butcher's Shop," and the "Janitor." We do not recommend the Goddard pictures as interpretation tests, as they are undoubtedly too easy. We do not think that the substitution of the colored pictures will have any effect on their standardization in Year III.

(2) Year IV—1: *Comparison of lines.* Terman gives, "Look closely and tell me which one is longer. Put your finger on the longest one." In our opinion "closely" will mean very little to the four year child and the use of both "longer" and "longest" may be confusing. In the four year child's language one line is the *long* one and the other is the *short* one. Terman comments, "As noted by Binet, success in this test depends on the comprehension of the verbal discussion, rather than on actual discrimination of length." This remark applies with almost equal force to a number of the other tests for the lower years.

(3) Year IV—5: *Comprehension of Questions.* In this test, as in several others, we do not think it necessary to impose any rigid time limit. In each case, however, we quote the time limit recommended by Terman. The truth of Terman's remark, that it probably requires more intelligence to describe what we ought to do in an imaginary situation than to do the right thing when the real situation is encountered, entirely depends on the complexity of the given situation. Descriptive and practical ability are too different to admit of comparison from the standpoint of the degree of intelligence involved. There is little doubt as to which kind of ability, however,

has the greatest social value. One of the criticisms that has already been offered is that the tests are too "literary," and favor the child with good language ability and little executive capacity in other directions.

(4) Year IV—6: *Repeating Four Digits.* We have avoided such expressions as "Go ahead" for "Begin;" "Clear on," for "So on," and "I am through" for "I have finished." These may be common in one community and not in others. Such expressions are, however, not only allowable but desirable, in the places where they are commonly used. It is, however, unwise to attempt to standardize colloquialisms for general use.

(5) Year V—2: *Naming Colors.* Terman declares himself against the question, "What color is that?" He states that it might bring the answer, "First color." The reply may be made that it certainly might, but it very rarely does. In any case our form of the question is such that it is made quite clear to the child that it is the name of the color that is required. "What color is that"—is almost universally the form used in everyday life.

(6) Year V—3: *Aesthetic Comparison.* For our use of "pretty" as against "prettiest," see note on Test IV—1.

(7) Year VI—2: *Missing Features.* Our form differs slightly from Terman's, but the difference is unessential. What is required in the test is made, we think, a little clearer by the question at the end.

(8) Year VII—2: *Description of Pictures.* Our procedure differs slightly from Terman's. We avoid the suggestion of enumeration by our form "Tell me all about this picture?" and at the same time give the needful prompting to full description, which some children require. We recommend Terman's pictures for this test, Goddard's being too simple.

(9) Year VII—5: *Differences between Common Objects.* The word "difference" is not always understood by the child. "Tell me in what way (or how) a butterfly is not like a fly?" is more nearly the child's form of expression.

(10) Year VII—Alt. 2. *Tying a Bow Knot.* We recommend tying of the bow knot by the child around a pencil or stick held by the examiner, as some children are apparently disconcerted by having to tie it around the examiner's finger.

(11) Year VIII—1: *Ball and Field.* This is one of the tests added to the Scale by Terman. Having applied it to several thousand children, we feel justified in altering slightly his procedure on the basis of our experience. For the eight year child part of the instructions are not very comprehensible, although suited to year twelve, to whom, also, it is applied. We refer particularly to the words "with....force" and "direction," in the explanation of the tests. We further believe that the intelligibility of the test has been improved by suggesting that the field is a very large one. Procedure

in looking for a ball in a small field may differ altogether from that pursued in a large area. In saying, "We don't know where it is or anything about it," we convey more clearly the fact that it may be in any part of the field than does Terman's remark about force and direction. We say, "Go about looking for the ball," which is free from the suggestion implied in "Go around," (which Terman objects to), and yet conveys the idea of continued search. Finally, the scoring of results is made much simpler by our injunction to the child not to lift the pencil. Terman publishes a score card of satisfactory and unsatisfactory responses. This scoring appears in several instances to be formulated upon an arbitrary principle, and with little reference to the intelligence of the response. Amongst the "superior" plans are given a series of concentric circles, with no connection between one and the other. If the child is supposed to have marked the path actually taken, he has missed the point that the path of search must be an unbroken line. Why this is superior to Number 8 on the score-card—a series of concentric rings all connecting at the gate—or to Number 10—a series of irregularly drawn rings roughly parallel—is not apparent; as a matter of fact, the course that would actually be pursued in a large field would be nearer Number 10 than any other of the concentric circle designs. Because there is no continuous line, Number 15—wheel spokes—and Number 19—fan shape—are to our thinking unsatisfactory, and indeed impossible. On the other hand the child who has made plan Number 21 may have a far more intelligent viewpoint of the problem than the child who does not realize that to get from one place to another he must have a connecting path. Either Numbers 14, 15, 16, 17, 18, 19, and 20 should be scored as failures or else our injunction, "Do not lift the pencil from the paper once you start," should be adopted. Very frequently a child will gain twelve year credit in this test and have very few, if any, intermediate successes between this test and Year XII.

Terman states that the problem may be regarded as a test of practical judgment, and unlike the majority of the other tests gives the subject a chance to show how he can meet the demands of a real, rather than an imagined, situation. On the other hand, it appears to us a problem which presents essentially an artificial and not a real situation. In a real situation of this kind, whilst a plan of search is pursued, it may seldom take the form of a geometrical measuring off of the ground, because facts such as the configuration of the surface, the size of the field, the amount of ground covered, and even "hunches" as to the likely places in which to look often influence the intelligent person's behavior. Many persons will proceed at once to the centre of the field and zigzag irregularly around it, and perhaps change the plan altogether by proceeding around the fence. Only after trial

and error method has proved fruitless would an intelligent person adopt a system of search. As the position of the ball is governed wholly by chance, the person adopting the trial and error method will very often happen upon the ball much sooner than the person who adopts a plan. Undoubtedly the tacit recognition of this fact would influence many to adopt a not too systematized method.

(12) Year VIII—2: *Counting Backwards.* Terman extends the time limit from twenty to forty seconds on the ground that intelligent and deliberate action often go together. The same reason may with equal justice be advanced against imposing a too rigid time limit in other tests.

(13) Year VIII—5: *Definitions Superior to Use.* Terman states that supplementary questions are never in order, and includes among satisfactory responses definitions which tell what class the object belongs to, although admitting that these definitions are of a particularly uneven value. Apparently, then, he accepts "A soldier is a person," or "A soldier is a man," whilst he would fail "A soldier fights," or "A soldier shoots." This hardly seems justified, since the child with the apperception of the soldier as shooting or fighting or marching has a clearer apperception of the idea than the child who merely says, "The soldier is a man." Again, with regard to a football, he passes, "It is a thing you play with," and fails "What they play with." The former definition does not tell any more about a football than the latter. One merely indicates greater facility in expression than the other. If the test is in any degree a test of apperception we cannot see why, when the idea is the same, crudity in expression should be penalized. Personally we prefer to regard all such responses as doubtful and would supplement the test by the suggestive remark, "But you have not told me enough about it."

(14) Year VIII—6: *Vocabulary Test.* Terman notes with regard to this valuable test, that "In a large majority of cases the Vocabulary Test alone will give us an intelligence quotient within ten percent of that secured by the entire Scale." This remark provides a very interesting comment on the objection that the Binet Scale as a whole is far more a test of comprehension of language and facility of oral expression than of practical ability in other directions.

(15) IX—2: *Arranging Weights.* Terman says of this test that "It tests the subject's intelligence in the manipulation of things, rather than his capacity for dealing with abstractions. It tests his ability to do something rather than his ability to express himself in language. It throws light upon certain factors of motor adaptation and practical judgment, which play a great part in the everyday life of the average human being."

In view of the importance of practical judgment, it is strange that better

provision has not been made in the Scale for the evaluation of this capacity. The chief objection to the Arranging Weights Test, as a measure of the individual's ability to adapt himself to a practical task, is that the test is not self-corrective in the sense that the subject does not realize his own error, and hence will not see any reason to correct a faulty method. When the blocks are arranged it is not, of course, apparent to the child whether he has them in the right order or not. The subject's realization of non-success is one of the chief factors in proceeding by experience in readjusting methods. Secondly, this test does not appear to be suitable for close standardization, as it appears in only two years of the Scale, years five and nine.

(16) IX—3: *Making Change.* We have preferred to keep to the Goddard procedure of using the actual coins to make change, as this is a much more practical task.

(17) IX—5: *Sentence Building.* The Stanford explanation of a sentence as being made up of words which say something is not very illuminating. We prefer to explain the word "sentence" as "a little story."

(18) X—4: *Reading and Report.* Terman in his discussion (see Page 264) defends the retention of this test. We agree that as a test of memory for ideas, especially when the subject is not warned that he is expected to recall the facts read, it is a valuable test. It measures not only reading ability, but also the depth of the impression or the understanding of ideas met with in a casual way. This depends very largely on the comprehension of the subject matter. As Terman remarks, defectives often read glibly but with little understanding. On the other hand, we certainly cannot agree that "Language growth, in fact, mirrors the entire mental development." Experience with defectives shows that individuals of equal intelligence—per Binet standards—have frequently very unequal development of language. In many cases of mental disorder, involving degenerative changes, there is often no diminution of language mastery, sometimes quite the reverse. The ideas are confused and flighty whilst the power of expression remains unimpaired. How then can the language development mirror the entire mental growth, if intelligence may increase or decrease, independent of the power of oral expression? The truth is that language mastery in relation to intelligence varies almost as much as any other specific ability. Our objection is not to language tests as a measure of intelligence, but to an excessive number of them in the Scale. Among the questions in the next test we find, "Why do we judge a person more by his actions than by his words?" A scale over-weighted with language tests is tending to reverse the usual methods of judgment.

(19) XII—7: *Interpretation of Pictures.* We include the injunction, "Explain this picture," in our first instructions. Since it is allowable to use this formula in case of failure, it seems more reasonable to give it to all children.

(20) XII—8: *Giving Similarities.* The test would be more justifiable from the standpoint of intelligibility if its form were, "A snake, a cow, and a sparrow are three animals that are not at all like each other, but there is something which is true of all of them. What is it?" Or, "A book, a teacher, and a newspaper are three very different things, but there is one thing that is true to say about all of them. What is it?"

References

Terman, Lewis M.—The Measurement of Intelligence. Houghton Mifflin Co., Boston, 1916.

Gillingham, Anna—The Bright Child and the School. *Journal of Educational Psychology,* May-June, 1919.

Doll, Edgar A.—A Brief Binet-Simon Scale. *The Psychological Clinic,* December, 1917, and January, 1918.

Goddard, H. H.—Two Thousand Normal Children Measured by the Binet Measuring Scale of Intelligence. *Pedagogical Seminary,* June, 1911.

Yerkes, Robert M.; Bridges, James W., and Hardwick, Rose S.—A Point Scale for Measuring Mental Ability. Warwick and York, Baltimore, 1915.

Pintner, Rudolf, and Paterson, Donald G.—A Scale of Performance Tests. D. Appleton & Co., New York, 1917.

Woolley, Helen Thompson—A Scale of Mental and Physical Measurements for Adolescents and Some of Its Uses. *Journal of Educational Psychology,* November, 1915.

Acknowledgements

The first part of this paper has been prepared with the assistance of Miss Bassett, Research Fellow, and Misses Flowers, Babcock, and Miller, members of the laboratory staff of the Research Department of The Training School, Vineland. Part II has been arranged in conjunction with Miss Hill.

It will be obvious to readers that this modification of the Binet-Simon Scale is based, as regards standardization, upon the Stanford Revision. The author gratefully acknowledges his great indebtedness to Professor Terman and his coworkers, whose work has met with such widespread approbation. This paper, therefore, makes no claims for originality as regards the tests themselves. It is hoped, however, that the provision of percentile tables, the selection of the tests of greater diagnostic value, and their general arrangement will result in greater convenience of application and better interpretation of the Scale than heretofore.